Available online at www.sciencedirect.com
SCIENCE DIRECT®
Computers and Electronics in Agriculture 43 (2004) 179–195
Computers and electronics in agriculture
www.elsevier.com/locate/compag

A guidance directrix approach to vision-based vehicle guidance systems

S. Han [a,*], Q. Zhang [b], B. Ni [c], J.F. Reid [d]

[a] *John Deere Ag Management Solutions, 4140 114th Street, Urbandale, IA 50322, USA*
[b] *Department of Agricultural Engineering, University of Illinois, Urbana, IL 61801, USA*
[c] *CNH Global, Burr Ridge Operations, Burr Ridge, IL 60527, USA*
[d] *John Deere Technology Center, Moline, IL 61265, USA*

Received 15 July 2003; received in revised form 18 December 2003; accepted 18 January 2004

Abstract

In a vision-based automatic agricultural vehicle guidance system for row-crop applications, finding guidance information from crop row structure is the key in achieving accurate control of the vehicle. This paper describes a robust procedure to obtain a guidance directrix. The procedure includes row segmentation by *K*-means clustering algorithm, row detection by a moment algorithm, and guidance line selection by a cost function. Auxiliary information, such as the known crop row spacing, is used to aid in the development of the guidance directrix. Two image data sets, one taken from a soybean field and the other taken from a corn field, were used to evaluate the accuracy of the proposed image processing procedure. The average RMS offset error from 30 soybean images was 1.0 cm with an average cost of 4.99. In contrast, the average RMS offset error from 15 corn images was 2.4 cm with an average cost of 7.27. The proposed image processing procedure was implemented on a vision-based guidance tractor.

Keywords: Tractor; Automatic guidance; Machine vision; Image processing

1. Introduction

Potential benefits of automated agricultural vehicles include increased productivity, increased application accuracy, and enhanced operation safety. The first stage development in agricultural vehicle automation, automatic vehicle guidance, has been studied for many years. A number of innovations were explored as early as 1920s (Willrodt, 1924; Sissons,

* Corresponding author. Tel.: +1-515-331-4675; fax: +1-515-331-4641.
E-mail address: hanshufeng@johndeere.com (S. Han).

0168-1699/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.compag.2004.01.007

1939; Schafer and Young, 1979). The earlier guidance systems demonstrated their technical feasibility, but commercialization was unsuccessful mainly due to the high cost of those systems.

The rapid advancement in electronics, computers, and computing technologies has inspired renewed interests in the development of vehicle guidance systems. Various guidance technologies, including mechanical guidance, optical guidance, radio navigation, and ultrasonic guidance, have been investigated (Reid et al., 2000; Tillett, 1991). In general, any guidance technology will need to provide positioning information of the vehicle, referenced either in a global coordinate system or in a local coordinate system.

Since early 1990s, Global Positioning System (GPS) receivers have been widely used as global guidance sensors (Bell, 2000; Larsen et al., 1994; Yukumoto et al., 2000). GPS-based guidance technology can be used for many field operations such as tillage, planting, cultivating, and harvesting. It has the potential to achieve completely autonomous navigation. However, the high cost of precision GPS receivers is the major obstacle for their widespread use in agricultural vehicle navigation.

Machine vision technology can be used to automatically guide a vehicle when crop row structure is distinguishable in a field. Typical applications include guiding a tractor for row-crop cultivation, or guiding a combine for harvest operation. The guidance sensor, i.e., the camera, is a local sensor because only the relative location of the vehicle, with respect to the crop rows, can be determined. Machine vision guidance has the advantage of using local features to fine-tune the vehicle navigation course. It has the technological characteristics closely resembling those possessed by a human operator, and thus has great potential for implementation of a vehicle guidance system (Wilson, 2000).

In a vision-based vehicle guidance system, finding guidance information from crop row structure is the key in achieving accurate control of the vehicle. A number of image processing techniques have been investigated to find the guidance course (directrix) from row-crop images. As examples, Reid et al. (1985) developed a binary thresholding strategy using a Bayes classification technique to effectively and accurately segment crop canopy and soil background for cotton crop at different growth stages. Gerrish et al. (1985) concluded in their study that thresholded intensity images alone will not work in all cases, and they showed that the combination of noise filtering, edge detection, thresholding, and re-scaling was the most promise technique. Image analysis using Hough transform to find crop rows was reported in several studies (Marchant and Brivot, 1995; Marchant, 1996).

Only a few vision guidance systems have been successfully developed and tested in field trials. Billingsley and Schoenfisch (1997) reported a vision guidance system that is relatively insensitive to additional visual ‘noise’ from weeds, while tolerating the fading out of one or more rows in a barren patch of the field. They showed that their system is capable of maintaining an accuracy of 2 cm. Since 1997, the University of Illinois research team has been working with industry partners to develop an automatic vision guidance tractor. The work has resulted in several US patents that disclose the guidance system. This paper describes the procedure used in the guidance system for obtaining the guidance directrix from row-crop images. The main objective of the reported research was to develop an image processing procedure that can be implemented in a vision-based guidance tractor in real-time with adequate accuracy. Examples are given to illustrate the robustness of the proposed image processing procedure.

2. Overview of the vision-based guidance system

The vision-based guidance system developed at the University of Illinois includes a vision sensor, an image processor, a DGPS receiver, a navigation planner, a steering controller, and a steering actuator (Fig. 1). The vision sensor is a Cohu 2100 series monochrome CCD camera (Cohu, Inc., San Diego, California). The camera has a fixed gain with an NIR bandpass filter at 800 nm. It acquires a scene image in the front of the vehicle. The use of the NIR filter can enhance the intensity (brightness) of crop pixels in the scene image. The image processor processes the acquired image to obtain a guidance directrix. The guidance directrix, along with DGPS positional information and other tractor states, is utilized in the navigation planner to determine a navigation command, i.e., the commanding wheel angle. Based on the difference between the commanding and the measured wheel angles, the steering controller computes an implementing steering control signal and sends it to drive the steering actuator. The steering actuator is an electrohydraulic (E/H) steering system that controls the turning rate of vehicle wheels. After the vehicle moves a short distance, the next loop starts with the vision sensor acquiring a new image.

A closed-loop control strategy is used in the steering controller, with the measured wheel angle as the feedback information. The control loop is executed at a faster rate than image processing loop to achieve a better vehicle control. Several steering controllers, including PID controller, feed-forward PID (FPID) controller, and fuzzy logic (FL) controller, have been developed and implemented in the guidance system (Qiu et al., 1999; Wu et al., 1999; Zhang, 1999).

As a linkage between the image processor and the steering controller, the navigation planner plays a key role in the guidance system. Development of the navigation planner has been discussed in Han et al. (2001). The image processor in Fig. 1 is critical for the success

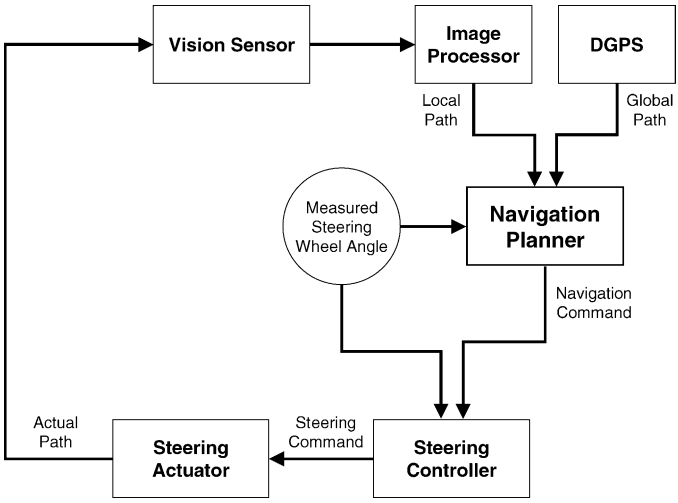


Fig. 1. Components of the vision-based guidance system developed at the University of Illinois.

of the vision-based guidance system. Development of the image processor to obtain the guidance directrix is discussed in the next section.

3. Development of the guidance directrix

To achieve high image processing speed, only a part of the entire scene image, called region-of-interest (ROI), is processed (Fig. 2). The ROI contains an upper region and a lower region, each includes two rows of crop to be tracked. A conventional *K*-means clustering algorithm is used to compute a threshold in the ROI. The threshold value depends on the pixel intensity distribution in the ROI and is used to identify pixels that represent crop or vegetation. Image segmentation is performed in four tracking windows within the ROI. A linear regression algorithm is applied in each tracking window to obtain a navigation line that represents the center of the crop row. Finally, all navigation lines in the image are combined to obtain a reliable guidance directrix. The procedure is described in the following paragraphs in details.

3.1. Row segmentation by K-means clustering

The image taken by the monochrome CCD camera is digitized using a frame grabber PXC200 (CyberOptics Corporation, Portland, Oregon) into a two-dimensional array of pixels, with each pixel represented by a gray level (GL) between 0 and 255. In a typical image scene (Fig. 2), the brighter pixels with higher GL represent crop, and the darker pixels with lower GL represent non-crop (e.g., soil). Image thresholding is the first step to convert the gray scale image into a binary image so that crop pixels are separated from the background.

If we construct a histogram of the pixel GL values, the pixels representing crop or vegetation would appear at the histogram's right side. The objective of image thresholding is to find a GL value (threshold) in the histogram such that any pixel with a higher GL value than that threshold would represent a crop pixel. A conventional *K*-means clustering algorithm is used to compute a threshold in the ROI.

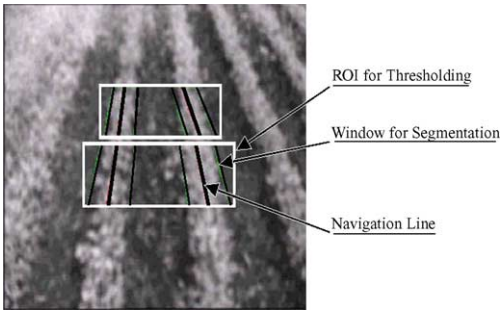


Fig. 2. Region-of-interest (ROI) for image thresholding, windows for image segmentation, and navigation lines for guidance.

K-means clustering is a partitioning method for grouping objects so that the within-group variance is minimized. By minimizing dissimilarity of each subset locally, the algorithm will globally yield an optimal dissimilarity of all subsets. The algorithm, as applied to image thresholding, is given by the following steps:

1. Initialize the (*K*) class centers. For simplicity, an equal-distance method is used to define the initial class centers:

$$\text{Center}_i^0 = \text{GL}_{\min} + \left(i - \frac{1}{2}\right)\frac{\text{GL}_{\max} - \text{GL}_{\min}}{K}, \quad i = 1, 2, ..., K \tag{1}$$

   where Center_i^0 is the initial class center for the *i*th class, $\text{GL}_{\max}$ and $\text{GL}_{\min}$ are the maximum and minimum GL values in the sample space, respectively.
2. Assign each point to its closest class center. The criterion to assign a point to a class is based on the Euclidean distance in the feature (GL) space using:

$$\text{Distance}_{i,j} = \text{abs}(\text{GL}_j - \text{Center}_i), \quad i = 1, 2, ..., K; \ \ j = 1, 2, ..., N \tag{2}$$

   where $\text{Distance}_{i,j}$ is the distance from the *j*th point to the *i*th class. *N* is the total number of points in the sample space.
3. Calculate the (*K*) new class centers from the mean of the points that are assigned to it. The new class centers are calculated by

$$\text{Center}_i^m = \frac{1}{N_i}\sum_{j=1}^{N_i}\text{GL}_j, \quad i = 1, 2, ...K \tag{3}$$

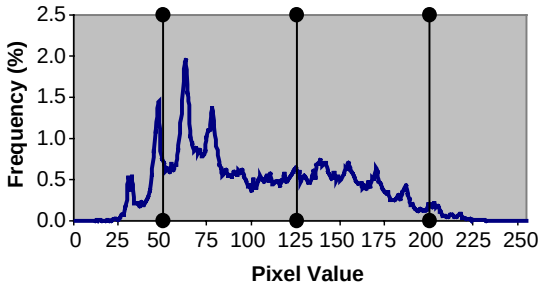
   where N_i is the total number of points that are assigned to the *i*th class in step 2.
4. Repeat from 2, if any class center has moved or if a threshold of movement has not been obtained. The criterion for measuring the change of the class centers is

$$\text{Measure} = \sqrt{\sum_{i=1}^{K}(\text{Center}_i^m - \text{Center}_i^{m-1})^2}, \quad i = 1, 2, ...K \tag{4}$$

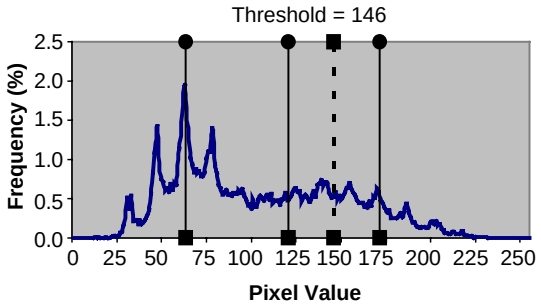
   where Center_i^m and Center_i^{m-1} are the new (or the current) *i*th class center and the old (or the previous) *i*th class center, respectively.
5. The threshold value is defined as the average of the *K*th class center and the $(K-1)$th class center:

$$\text{Threshold} = \tfrac{1}{2}(\text{Center}_K + \text{Center}_{K-1}) \tag{5}$$

As an example, the gray-scale image in Fig. 2 was sampled every pixel to obtain a histogram of the GL values (Fig. 3a). The minimum GL value was 13, and the maximum GL value was 238. Three classes were selected in this example ($K = 3$). The initial class centers were (50, 125, 200). After eight iterations as described above, the new class centers were found as (63, 120, 171) (Fig. 3b). The threshold was selected as 146 (Eq. (5)). Using this threshold value, a segmented binary image was obtained (Fig. 4).



(a). Initial K (K=3) classes.



(b). Class centers after segmentation.

Fig. 3. *K*-means clustering method to find a threshold.

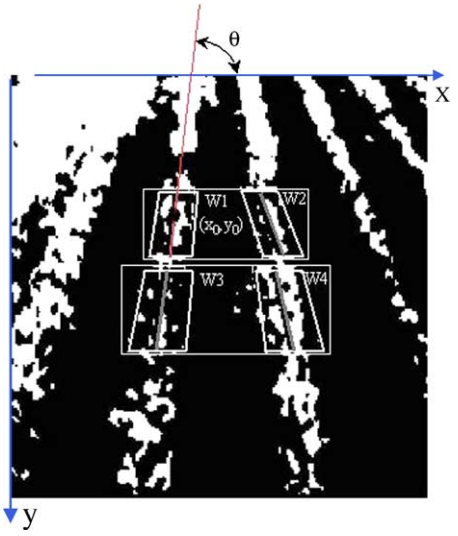


Fig. 4. Row detection from the segmented binary image.

3.2. Row detection by a moment algorithm

Four tracking windows within the ROI are selected, and each window keeps track of a segment of one crop row (Fig. 4). Within each tracking window, the crop row can be approximated by a straight line with a centroid point, (x_0, y_0), and an orientation angle, θ.

The location of the centroid point within a window is determined by

$$x_0 = \frac{\sum_{n=0}^{\Delta y}\sum_{m=x1(n)}^{x2(n)} m \times \mathrm{GL}(m,n)}{\sum_{n=0}^{\Delta y}\sum_{m=x1(n)}^{x2(n)} \mathrm{GL}(m,n)}$$
$$y_0 = \frac{\sum_{n=0}^{\Delta y}\sum_{m=x1(n)}^{x2(n)} n \times \mathrm{GL}(m,n)}{\sum_{n=0}^{\Delta y}\sum_{m=x1(n)}^{x2(n)} \mathrm{GL}(m,n)} \tag{6}$$

where $\mathrm{GL}(m,n)$ is the GL value of the vegetation pixel at location (m, n) in the binary image, Δy is the height (in pixel) of the tracking window, $x1(n)$ and $x2(n)$ are the left and right interception points of the tracking window with a horizontal line $y = n$, respectively. Note that the operation is performed inside the tracking window with background pixels excluded. Since the tracking window is an irregular rectangle, both $x1(n)$ and $x2(n)$ depend on the index n.

The orientation angle, θ, is determined by a moment algorithm:

$$\theta = \tan^{-1}\left\{\frac{-(M_{20}-M_{02}) \pm \sqrt{(M_{20}-M_{02})^2 - 4M_{11}}}{2M_{11}}\right\} \tag{7}$$

where

$$M_{20} = \sum_{n=0}^{\Delta y}\sum_{m=x1(n)}^{x2(n)} (m-x_0)^2 \times \mathrm{GL}(m,n)$$
$$M_{02} = \sum_{n=0}^{\Delta y}\sum_{m=x1(n)}^{x2(n)} (n-y_0)^2 \times \mathrm{GL}(m,n) \tag{8}$$
$$M_{11} = \sum_{n=0}^{\Delta y}\sum_{m=x1(n)}^{x2(n)} (m-x_0)(n-y_0) \times \mathrm{GL}(m,n)$$

Eqs. (6)–(8) are applied to each of the four tracking windows to obtain four navigation lines. The combination of these navigation lines determines a guidance directrix.

3.3. Navigation line selection by a cost function

The quality of a navigation line is measured by a cost function defined below:

$$\cos t = \frac{M_{20}}{\sum_{n=0}^{\Delta y}\sum_{m=x1(n)}^{x2(n)} \mathrm{GL}(m,n)} \tag{9}$$

where M_{20} is defined in Eq. (8).

The cost function is essentially a second moment about the y-axis. Greater scattering of vegetation pixels results in a higher cost value. Conversely, less scatter (a better quality navigation line) results in a lower cost value. Cost functions are applied to the upper and lower regions in the ROI (Fig. 2) to determine whether the navigation lines from Eqs. (6)–(8) are directly accepted, or need to be further processed by the following criteria:

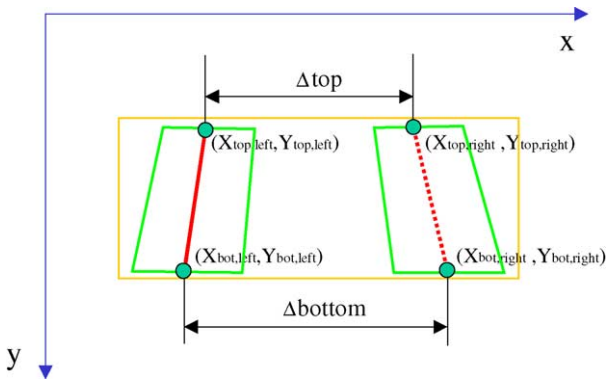


Fig. 5. Calculated navigation line from the known row spacing.

1. If the costs of the navigation line in both the left and the right windows are acceptable, both navigation lines are accepted.
2. If the cost of the navigation line in the left window is acceptable and the cost of the navigation line in the right window is not acceptable, the left navigation line is accepted and the right navigation line is calculated.
3. If the cost of the navigation line in the right window is acceptable and the cost of the navigation line in the left window is not acceptable, the right navigation line is accepted and the left navigation line is calculated.
4. If the costs of the navigation line in both the left and the right windows are not acceptable, the navigation lines determined from the previous image(s) are used.

The cost threshold to accept or reject a navigation line was determined a priori by visual examination. Typically, a cost of 3 or less indicates a good navigation line, but a navigation line with a cost between 3 and 8 is also acceptable. When the cost value is above 9, the navigation line is considered as unacceptable.

The top points on navigation lines in the left and the right windows should be separated by a known crop row spacing. The same should be true for the two bottom points on those navigation lines. Mathematically, these conditions are expressed as

$$\begin{aligned} x_{\text{top,right}} &= x_{\text{top,left}} + \Delta_{\text{top}} \\ y_{\text{top,right}} &= y_{\text{top,left}} \\ x_{\text{bot,right}} &= x_{\text{bot,left}} + \Delta_{\text{bottom}} \\ y_{\text{bot,right}} &= y_{\text{bot,left}} \end{aligned} \tag{10}$$

where Δ_{top} and Δ_{bottom} are the distances between two crop rows in the image space (Fig. 5). Because of the perspective view of the scene, Δ_{top} and Δ_{bottom} are not the same in the image space. Eq. (10) is used to obtain the two end points on the calculated navigation line under condition (2), (3), or (4) above. The centroid point, (x_0, y_0), and the orientation angle, θ, of the calculated navigation line can be further obtained from Eq. (10).

Many factors, such as poor crop development, bare spots, or shadows, often make it difficult to recognize crop rows correctly in one or more tracking windows. Using a cost

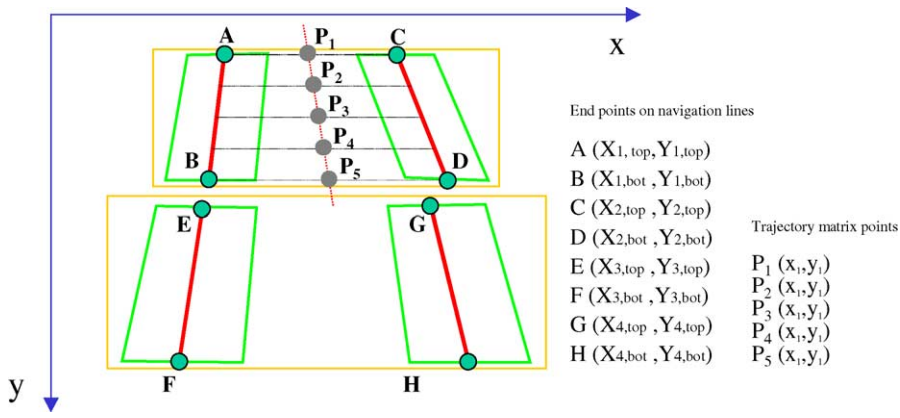


Fig. 6. Relative positions of navigation lines in the tracking windows.

function is an effective way to identify poor-quality navigation lines in tracking windows, and to replace these lines with calculated lines using auxiliary information—a known crop row spacing in this case. It is obvious that using multiple tracking windows can improve the robustness of the image processing.

3.4. Improving row tracking reliability

In addition to the cost function, another quality measure for navigation lines in tracking windows is the relative positions among these lines. As shown in Fig. 6, each navigation line can be represented by two end points, $(X_{i,\text{top}}, Y_{i,\text{top}})$ and $(X_{i,\text{bot}}, Y_{i,\text{bot}})$, $i = 1, 2, 3, 4$. In an ideal situation, the distances between A and C, B and D, E and G, F and H, when transformed to the vehicle (ground) coordinate system, should be close to a known row spacing. Due to image processing error, relative positions among the navigation lines may not be perfect. The following criteria are used to adjust navigation lines:

1. All the tracking windows are good, i.e.,

$$\begin{aligned}
&\{T(X_{2,\text{top}}) - T(X_{1,\text{top}})\} > S_{\min} \quad \text{and} \quad \{T(X_{2,\text{top}}) - T(X_{1,\text{top}})\} < S_{\max}\\
&\{T(X_{2,\text{bot}}) - T(X_{1,\text{bot}})\} > S_{\min} \quad \text{and} \quad \{T(X_{2,\text{bot}}) - T(X_{1,\text{bot}})\} < S_{\max}\\
&\{T(X_{4,\text{top}}) - T(X_{3,\text{top}})\} > S_{\min} \quad \text{and} \quad \{T(X_{4,\text{top}}) - T(X_{3,\text{top}})\} < S_{\max}\\
&\{T(X_{4,\text{bot}}) - T(X_{3,\text{bot}})\} > S_{\min} \quad \text{and} \quad \{T(X_{4,\text{bot}}) - T(X_{3,\text{bot}})\} < S_{\max}\\
&\{T(X_{2,\text{top}}) - T(X_{3,\text{top}})\} > S_{\min} \quad \text{and} \quad \{T(X_{2,\text{top}}) - T(X_{3,\text{top}})\} < S_{\max}\\
&\{T(X_{2,\text{bot}}) - T(X_{3,\text{bot}})\} > S_{\min} \quad \text{and} \quad \{T(X_{2,\text{bot}}) - T(X_{3,\text{bot}})\} < S_{\max}\\
&\{T(X_{4,\text{top}}) - T(X_{1,\text{top}})\} > S_{\min} \quad \text{and} \quad \{T(X_{4,\text{top}}) - T(X_{1,\text{top}})\} < S_{\max}\\
&\{T(X_{4,\text{bot}}) - T(X_{1,\text{bot}})\} > S_{\min} \quad \text{and} \quad \{T(X_{4,\text{bot}}) - T(X_{1,\text{bot}})\} < S_{\max}
\end{aligned} \tag{11}$$

where $T\{\ \}$ is the coordinate transformation from image space to vehicle space, $S_{\min}$ is the minimum acceptable row spacing, $S_{\max}$ is the maximum acceptable row spacing. $S_{\min} = (S_0 - \Delta S)$ and $S_{\max} = (S_0 + \Delta S)$. S_0 is the ideal row spacing, and ΔS is the

acceptable error in row spacing. No adjustment is needed for the navigation lines in this case.

2. Top-left tracking window is no good, i.e.,

$$\begin{aligned}
&\{T(X_{2,\text{top}}) - T(X_{1,\text{top}})\} < S_{\min} \quad \text{or} \quad \{T(X_{2,\text{top}}) - T(X_{1,\text{top}})\} > S_{\max} \\
&\{T(X_{2,\text{bot}}) - T(X_{1,\text{bot}})\} < S_{\min} \quad \text{or} \quad \{T(X_{2,\text{bot}}) - T(X_{1,\text{bot}})\} > S_{\max} \\
&\{T(X_{4,\text{top}}) - T(X_{3,\text{top}})\} > S_{\min} \quad \text{and} \quad \{T(X_{4,\text{top}}) - T(X_{3,\text{top}})\} < S_{\max} \\
&\{T(X_{4,\text{bot}}) - T(X_{3,\text{bot}})\} > S_{\min} \quad \text{and} \quad \{T(X_{4,\text{bot}}) - T(X_{3,\text{bot}})\} < S_{\max} \\
&\{T(X_{2,\text{top}}) - T(X_{3,\text{top}})\} > S_{\min} \quad \text{and} \quad \{T(X_{2,\text{top}}) - T(X_{3,\text{top}})\} < S_{\max} \\
&\{T(X_{2,\text{bot}}) - T(X_{3,\text{bot}})\} > S_{\min} \quad \text{and} \quad \{T(X_{2,\text{bot}}) - T(X_{3,\text{bot}})\} < S_{\max}
\end{aligned} \tag{12}$$

The navigation line in the top-left tracking window is reset by Eq. (10).

3. Top-right tracking window is no good, i.e.,

$$\begin{aligned}
&\{T(X_{2,\text{top}}) - T(X_{1,\text{top}})\} < S_{\min} \quad \text{or} \quad \{T(X_{2,\text{top}}) - T(X_{1,\text{top}})\} > S_{\max} \\
&\{T(X_{2,\text{bot}}) - T(X_{1,\text{bot}})\} < S_{\min} \quad \text{or} \quad \{T(X_{2,\text{bot}}) - T(X_{1,\text{bot}})\} > S_{\max} \\
&\{T(X_{4,\text{top}}) - T(X_{3,\text{top}})\} > S_{\min} \quad \text{and} \quad \{T(X_{4,\text{top}}) - T(X_{3,\text{top}})\} < S_{\max} \\
&\{T(X_{4,\text{bot}}) - T(X_{3,\text{bot}})\} > S_{\min} \quad \text{and} \quad \{T(X_{4,\text{bot}}) - T(X_{3,\text{bot}})\} < S_{\max} \\
&\{T(X_{4,\text{top}}) - T(X_{1,\text{top}})\} > S_{\min} \quad \text{and} \quad \{T(X_{4,\text{top}}) - T(X_{1,\text{top}})\} < S_{\max} \\
&\{T(X_{4,\text{bot}}) - T(X_{1,\text{bot}})\} > S_{\min} \quad \text{and} \quad \{T(X_{4,\text{bot}}) - T(X_{1,\text{bot}})\} < S_{\max}
\end{aligned} \tag{13}$$

The navigation line in the top-right tracking window is reset by Eq. (10).

4. Bottom-left tracking window is no good, i.e.,

$$\begin{aligned}
&\{T(X_{2,\text{top}}) - T(X_{1,\text{top}})\} > S_{\min} \quad \text{and} \quad \{T(X_{2,\text{top}}) - T(X_{1,\text{top}})\} < S_{\max} \\
&\{T(X_{2,\text{bot}}) - T(X_{1,\text{bot}})\} > S_{\min} \quad \text{and} \quad \{T(X_{2,\text{bot}}) - T(X_{1,\text{bot}})\} < S_{\max} \\
&\{T(X_{4,\text{top}}) - T(X_{3,\text{top}})\} < S_{\min} \quad \text{or} \quad \{T(X_{4,\text{top}}) - T(X_{3,\text{top}})\} > S_{\max} \\
&\{T(X_{4,\text{bot}}) - T(X_{3,\text{bot}})\} < S_{\min} \quad \text{or} \quad \{T(X_{4,\text{bot}}) - T(X_{3,\text{bot}})\} > S_{\max} \\
&\{T(X_{4,\text{top}}) - T(X_{1,\text{top}})\} > S_{\min} \quad \text{and} \quad \{T(X_{4,\text{top}}) - T(X_{1,\text{top}})\} < S_{\max} \\
&\{T(X_{4,\text{bot}}) - T(X_{1,\text{bot}})\} > S_{\min} \quad \text{and} \quad \{T(X_{4,\text{bot}}) - T(X_{1,\text{bot}})\} < S_{\max}
\end{aligned} \tag{14}$$

The navigation line in the bottom-left tracking window is reset by Eq. (10).

5. Bottom-right tracking window is no good, i.e.,

$$\begin{aligned}
&\{T(X_{2,\text{top}}) - T(X_{1,\text{top}})\} > S_{\min} \quad \text{and} \quad \{T(X_{2,\text{top}}) - T(X_{1,\text{top}})\} < S_{\max} \\
&\{T(X_{2,\text{bot}}) - T(X_{1,\text{bot}})\} > S_{\min} \quad \text{and} \quad \{T(X_{2,\text{bot}}) - T(X_{1,\text{bot}})\} < S_{\max} \\
&\{T(X_{4,\text{top}}) - T(X_{3,\text{top}})\} < S_{\min} \quad \text{or} \quad \{T(X_{4,\text{top}}) - T(X_{3,\text{top}})\} > S_{\max} \\
&\{T(X_{4,\text{bot}}) - T(X_{3,\text{bot}})\} < S_{\min} \quad \text{or} \quad \{T(X_{4,\text{bot}}) - T(X_{3,\text{bot}})\} > S_{\max} \\
&\{T(X_{2,\text{top}}) - T(X_{3,\text{top}})\} > S_{\min} \quad \text{and} \quad \{T(X_{2,\text{top}}) - T(X_{3,\text{top}})\} < S_{\max} \\
&\{T(X_{2,\text{bot}}) - T(X_{3,\text{bot}})\} > S_{\min} \quad \text{and} \quad \{T(X_{2,\text{bot}}) - T(X_{3,\text{bot}})\} < S_{\max}
\end{aligned} \tag{15}$$

The navigation line in the bottom-right tracking window is reset by Eq. (10).

3.5. Obtaining a guidance directrix

The guidance directrix in the image coordinate system (ICS) can be presented by an image trajectory matrix as (Han et al., 2001):

$$T^I_{t_k} = \begin{bmatrix} P^I_{1,t_k} \\ \vdots \\ P^I_{i,t_k} \\ \vdots \\ P^I_{m,t_k} \end{bmatrix} = \begin{bmatrix} x^I_{1,t_k} & y^I_{1,t_k} \\ \vdots & \vdots \\ x^I_{i,t_k} & y^I_{i,t_k} \\ \vdots & \vdots \\ x^I_{m,t_k} & y^I_{m,t_k} \end{bmatrix} \tag{16}$$

where $T^I_{t_k}$ is the image trajectory matrix at time $t = t_k$; $P^I_{i,t_k} = [x^I_{i,t_k} y^I_{i,t_k}]$ is the position of the ith point on the guidance directrix; and m is the total number of points on the guidance directrix.

For simplicity, five ($m = 5$) points were selected and included in the image trajectory matrix from the following equation (Fig. 6):

$$\begin{aligned} [x_1, y_1] &= [\tfrac{1}{2}(X_{1,\text{top}} + X_{2,\text{top}}), Y_{1,\text{top}}] \\ [x_5, y_5] &= [\tfrac{1}{2}(X_{1,\text{bot}} + X_{2,\text{bot}}), Y_{1,\text{bot}}] \\ [x_3, y_3] &= [\tfrac{1}{2}(x_1 + x_5), \tfrac{1}{2}(y_1 + y_5)] \\ [x_2, y_2] &= [\tfrac{1}{2}(x_1 + x_3), \tfrac{1}{2}(y_1 + y_3)] \\ [x_4, y_4] &= [\tfrac{1}{2}(x_3 + x_5), \tfrac{1}{2}(y_3 + y_5)] \end{aligned} \tag{17}$$

Only two navigation lines in the upper region were used for guidance. The navigation lines in the lower region were primarily used for improving the tracking reliability.

4. Results and discussion

In order to evaluate the performance of the image processing procedure under outdoor conditions, the vision sensor (camera) was mounted at the front of the tractor and was tilted about 30° from normal in the direction of travel. A 6 mm lens was used for the camera. A 300-MHz Pentium II computer operating under Windows® 98 was installed in the tractor cab for image acquisition and processing. In the following paragraphs, the quantitative analysis of the image processing procedure is provided using two sets of test images.

4.1. Images from soybean crop

A series of 30 images were taken in a soybean field when the crop was about 8 cm (3 in.) tall at a V3 development stage. The images were acquired under bright sunlight conditions. The soybean crop had a continuous row structure at this stage. Visually, the crop rows can be easily separated from the soil background. Fig. 7a1 and b1 shows two example images from this data set.

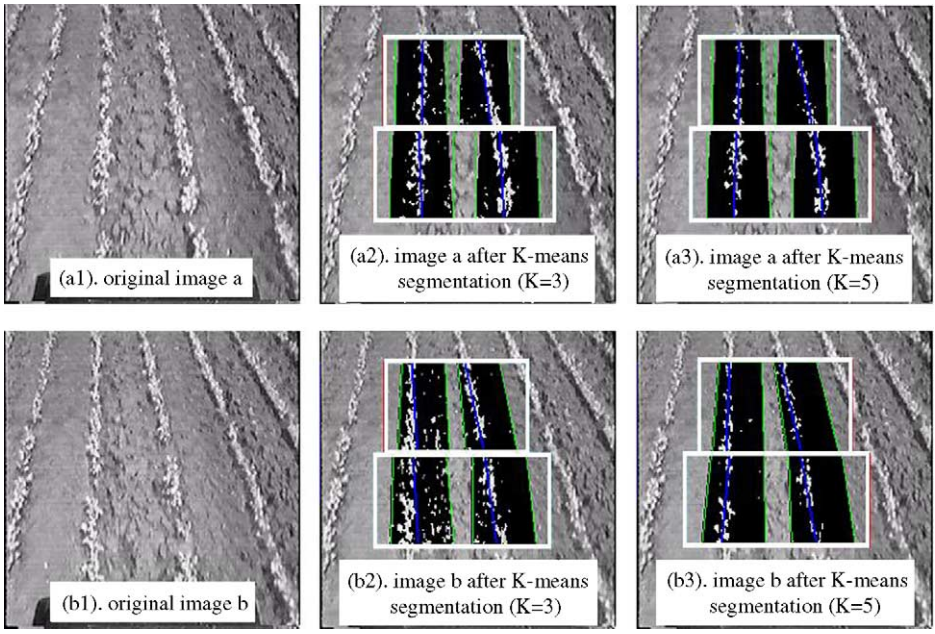


Fig. 7. Two example soybean images: (a1) and (b1) are the original images; (a2) and (b2), (a3) and (b3) show the results after *K*-means segmentation when $K = 3$ and 5, respectively.

The image size from the frame grabber was 256 × 243 ($W \times H$). In order to include two (and only two) crop rows in the ROI, the tracking window sizes were set to 55 × 75 for the bottom two windows and 45 × 75 for the top two windows (Fig. 7). Trajectory matrices for both the upper region and the lower region were calculated by the procedure described in Section 3.5. The two matrices were compared with the reference trajectory matrices obtained by manually (visually) selecting the crop rows from the image. An offset error for any point in the trajectory matrix, (x_i, y_i), is defined by

$$E_i = x_i - x_i^0 \tag{18}$$

where x_i^0 is the corresponding *x*-coordinate in the reference trajectory matrix at location $y = y_i$. Three points in each of the two trajectory matrices, i.e., the top, middle, and bottom points (Eq. (17)), were selected to quantify the offset error of an entire image.

A cost was calculated by Eq. (9) for each tracking window of an image. The average of four costs associated with four tracking windows of an image was used to quantify the cost of the entire image. The offset errors and costs for each soybean image are shown in Table 1. The average RMS offset error for this data set is 0.93, which translates to a ground distance of 1.0 cm in the center of the image. The average cost for this data set is 4.99.

It was found that the number of classes, *K*, in the *K*-mean clustering algorithm, had a significant impact on the performance of the proposed procedure. Fig. 8 shows the relationship between the average RMS offset error (for the entire data set) and the number of classes, *K*, in the *K*-means algorithm. The RMS offset error was reduced from 3.17 when $K = 3$ to

Table 1
Offset error and costs for soybean images ($K = 5$)

Image ID	Offset error (pixel)		Costs	
	RMS	Maximum	Mean	Maximum
1	1.16	1.75	4.18	5.32
2	0.82	1.51	5.78	6.50
3	1.22	2.07	4.52	5.75
4	1.43	2.99	5.10	5.89
5	0.91	1.30	4.72	7.44
6	0.75	1.24	5.22	6.79
7	0.77	1.67	4.73	6.14
8	0.57	1.19	5.77	7.57
9	0.78	1.16	5.25	6.29
10	0.78	1.50	4.71	5.49
11	0.88	1.61	4.65	5.62
12	1.01	1.84	4.36	5.72
13	1.50	3.15	5.43	6.23
14	0.61	0.88	4.48	5.98
15	0.89	1.61	4.58	4.97
16	0.64	1.22	5.11	5.92
17	0.69	1.29	4.78	5.57
18	0.67	1.15	4.55	5.33
19	0.77	1.46	4.84	6.30
20	0.46	0.92	4.80	5.70
21	1.09	1.74	4.82	6.28
22	0.76	1.53	5.00	7.16
23	1.48	2.32	5.13	6.43
24	1.83	3.37	5.56	8.42
25	0.33	0.66	5.14	7.03
26	0.31	0.46	5.32	7.50
27	0.42	0.85	5.59	6.43
28	2.24	4.65	6.31	7.63
29	0.82	1.21	5.06	6.74
30	1.16	1.76	4.26	5.07
Average	0.93	1.67	4.99	6.31

0.81 when $K = 7$. Further increasing K caused too many crop pixels to be segmented out. Fig. 7a2 and a3, b2 and b3 illustrates the effects of K on the image processing quality.

4.2. Images from corn crop

A series of 15 images were taken in a corn field when the crop was about 20 cm (3 in.) tall at a V4 development stage. Again, the images were acquired under bright sunlight conditions. The crop was too small to form a continuous row structure at this stage. In addition, crop development was not uniform, resulting in small plants and/or no plants at many locations. Although crop rows were visually identifiable, the non-continuous row structure posed a great challenge for reliable image processing. Fig. 9a1 and b1 shows two example images from this data set.

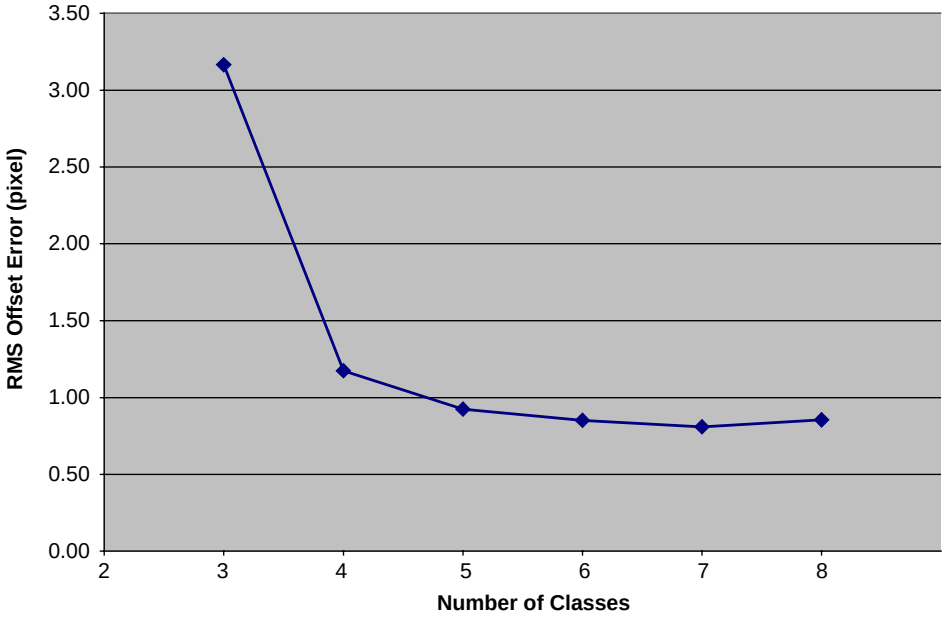


Fig. 8. Relationship between the average RMS offset error and the number of classes for soybean images.

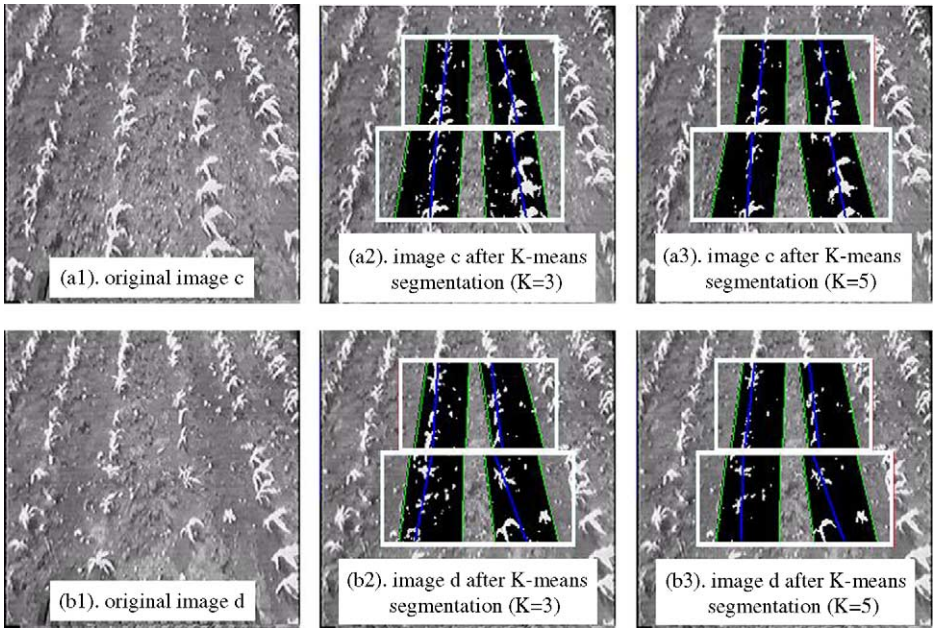


Fig. 9. Two example corn images: (a1) and (b1) are the original images; (a2) and (b2), (a3) and (b3) show the results after *K*-means segmentation when $K = 3$ and 5, respectively.

Table 2
Offset error and costs for corn images ($K = 5$)

Image ID	Offset error (pixel)		Costs	
	RMS	Maximum	Mean	Maximum
2	2.25	4.75	6.59	8.37
3	1.10	2.41	6.76	7.44
4	2.64	5.77	6.41	8.21
5	0.53	0.93	5.52	7.55
6	1.23	1.57	6.36	8.39
7	1.78	3.13	7.29	8.07
8	3.52	7.67	8.52	12.12
9	2.30	3.70	6.55	7.92
10	1.37	1.89	6.82	8.56
11	4.71	10.31	9.16	14.99
12	2.60	5.76	6.62	7.97
13	1.00	2.00	6.55	7.66
14	2.48	5.28	7.56	10.54
15	2.78	4.68	8.67	9.34
16	3.39	6.08	9.64	11.60
Average	2.25	4.40	7.27	9.25

The image size, tracking window sizes, and the evaluation process were the same as for the soybean data set, as discussed in the previous section. The offset errors and costs for each corn images are shown in Table 2. The average RMS offset error for this data set is 2.25, which translates to a ground distance of 2.4 cm in the center of the image. The average cost for this data set is 7.27. These numbers are much larger than for the soybean data set, indicating the difficulty in correctly identifying crop rows.

The relationship between the average RMS offset error (for the entire data set) and the number of classes, K, in the K-means algorithm is shown in Fig. 10. The RMS offset error was slightly reduced from 2.34 when $K = 3$ to 2.20 when $K = 6$. However, the effect of K on the RMS offset error was negligible for this data set. Fig. 9a2 and a3, b2 and b3 illustrates the effects of K on the image processing quality.

5. Conclusions and future work

The image processing procedure with guidance directrix approach achieved good accuracy from two test data sets: 1.0 cm average RMS offset error for a set of soybean images and 2.4 cm average RMS offset error for a set of corn images. The procedure is acceptable for real-time vision guidance applications in terms of its accuracy.

K-means algorithm in the procedure is the major limitation to the processing speed. Using the above-described hardware, images were collected and processed at a rate of 6–10 Hz, depending on image scene complexity. This processing speed is acceptable for real-time applications if the controller output rate is faster and is independent of the image update rate. With our implementation, the controller update rate is 50 Hz so the image processing speed was deemed acceptable.

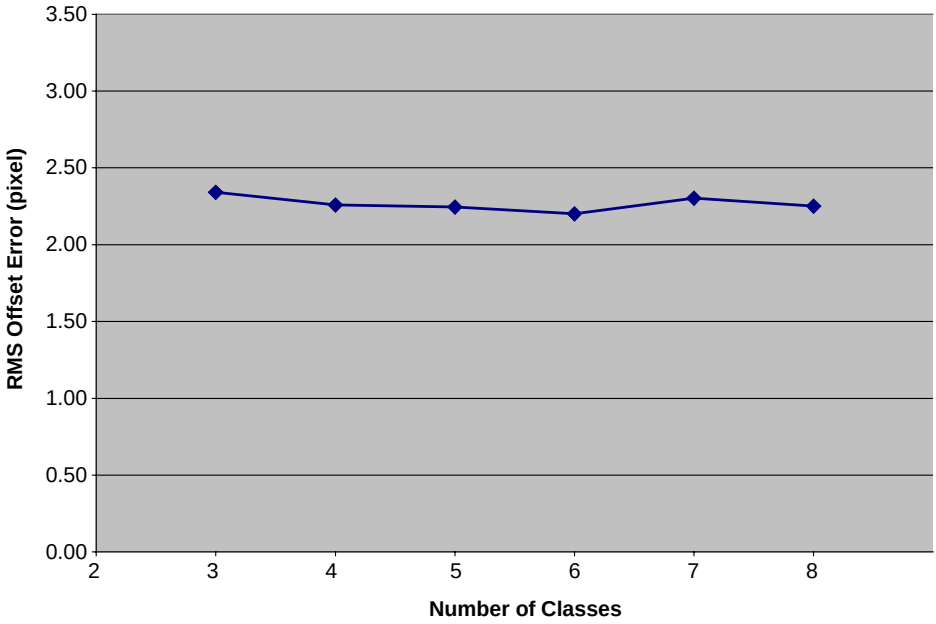


Fig. 10. Relationship between the average RMS offset error and the number of classes for corn images.

Although the results from the above two sets of test images were favorable, the performance of the proposed image processing procedure could degrade significantly under many adverse environmental conditions such as weeds or crop residue in the ROI. Quantitative evaluation of the algorithm under these conditions was not performed. However, using four tracking windows should increase the robustness of image processing since the navigation lines in one or two poor-quality tracking windows could be calculated from the lines in other good-quality tracking windows.

The proposed image processing procedure has been implemented on a vision-based guidance tractor. However, quantitative performance evaluation of the automatically guided tractor is yet to be done. The fusion of vision sensor with other navigation sensors, such as GPS, to provide more robust guidance directrix should be an active research topic in the future.

Acknowledgements

This material presented in this paper is based upon work supported by the Case Corporation, the University of Illinois at Urbana-Champaign (UIUC) Campus Research Board, and the United States Department of Agriculture (USDA) HATCH Funds (ILLU-10-0306, ILLU-10-0316). Any opinions, findings, and conclusions expressed in this publication are those of the authors and do not necessarily reflect the views of Case Corporation, UIUC, and USDA.

References

Bell, T., 2000. Automatic tractor guidance using carrier-phase differential GPS. Computers and Electronics in Agriculture 25, 53–66.

Billingsley, J., Schoenfisch, M., 1997. The successful development of a vision guidance system for agriculture. Computers and Electronics in Agriculture 16, 147–163.

Gerrish, J.B., Stockman, G.C., Mann, L., Hu, G., 1985. Image rocessing for path-finding in agricultural field operations. ASAE Paper #853037, ASAE, St. Joseph, MI, USA.

Han, S., Zhang, Q., Reid, J.F., 2001. A navigation planner for automatic tractor guidance using machine vision and DGPS. In: Proceedings of the 1st IFAC Conference on Telematics Applications in Automation and Robotics, TA2001. Weingarten, Germany, 24–26 July 2001, IFAC, Germany.

Larsen, W.E., Nielsen, G.A., Tyler, D.A., 1994. Precision navigation with GPS. Computers and Electronics in Agriculture 11, 85–95.

Marchant, J.A., 1996. Tracking of row structure in three crops using image analysis. Computers and Electronics in Agriculture 15, 161–179.

Marchant, J.A., Brivot, R., 1995. Real time tracking of plant rows using a Hough transform. Real Time Imaging 1, 363–375.

Qiu, H., Zhang, Q., Reid, J.F., Wu, D., 1999. Modeling and simulation of an electrohydraulic steering system. ASAE Paper #993076, ASAE, St. Joseph, MI, USA.

Reid, J.F., Searcy, S.W., Babowicz, R.J., 1985. Determining a guidance directrix in row crop images. ASAE Paper #85-3549, ASAE, St. Joseph, MI, USA.

Reid, J.F., Zhang, Q., Noguchi, N., Dickson, M., 2000. Agricultural automatic guidance research in North America. Computers and Electronics in Agriculture 25, 155–167.

Schafer, R.L., Young, R.E., 1979. An automatic guidance system for tractors. Transactions of the ASAE 22 (1), 46–49, 56.

Sissons, R., 1939. Plowing in circles saves time. The Prairie Farmer 111 (20), 7.

Tillett, N.D., 1991. Automatic guidance sensors for agricultural field machines: a review. Journal of Agricultural Engineering Research 50, 167–187.

Willrodt, F.L., 1924. Steering attachment for tractors. U.S. Patent No. 1506706.

Wilson, J.N., 2000. Guidance of agricultural vehicles—a historical perspective. Computers and Electronics in Agriculture 25, 3–9.

Wu, D., Zhang, Q., Reid, J.F., Qiu, H., 1999. Adaptive control of electrohydraulic steering system for wheel-type agricultural tractors. ASAE Paper #993079, ASAE, St. Joseph, MI, USA.

Yukumoto, O., Matsuo, Y., Noguchi, N., 2000. Robotization of agricultural vehicles (Part 2)—description of the tilling robot. Journal of Agricultural Engineering Research 34, 107–114.

Zhang, Q., 1999. Velocity control on a hydraulic system using feedforward–feedback control. ASAE Paper #991106, ASAE, St. Joseph, MI, USA.